



USAID
FROM THE AMERICAN PEOPLE

ተግባራዊ ጥራት ለህጻናት
ትምህርት ለግልጽ ልማት
ለ ግልጽ ልማት



المملكة المغربية
وزارة التربية الوطنية
والتكوين المهني

USAID/MOROCCO READING FOR SUCCESS-SMALL SCALE EXPERIMENTATION ACTIVITY

SUMMER ENRICHMENT PROGRAM: EARLY GRADE READING ASSESSMENT REPORT

Contract No. AID-608-TO-15-00002

This document was produced for review by the United States Agency for International Development. It was prepared by Chemonics International, Inc. and School-to-School International.

USAID/MOROCCO READING FOR SUCCESS-SMALL SCALE EXPERIMENTATION ACTIVITY

SUMMER ENRICHMENT PROGRAM: EARLY GRADE READING ASSESSMENT REPORT

CONTRACTED UNDER TASK ORDER NO. AID-608-TO-15-00002

DISCLAIMER

The authors' views expressed in this document do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

ACKNOWLEDGEMENTS

The School-to-School International and Chemonics International team would like to thank USAID/Morocco for their support. Particularly, we wish to acknowledge the important contribution of Thomas LeBlanc and Mariam Britel-Swift from the education office at USAID/Morocco. The USAID/Morocco education office has provided guidance and support throughout the preparation and administration of the study.

We would also like to recognize the staff from the Ministry of Education and Vocational Training and the National Center for Evaluations and Exams led by Mohammed Sassi, director of the center; Ahmed Chaibi, head of the evaluation division; Said Bouderga, head of the examinations division; and Hafida Kach, manager at the National Center for Evaluations and Exams. The National Center for Evaluations and Exams has played an important role at all stages of the study: sampling, tool development, identifying data collectors, development of the data analysis plan, and validation of the findings. We would also like to extend another special thank you to all the enumerators and data collectors who served as orientation inspectors and counselors. Their commitment, professionalism, and test-administration expertise played a key role in making the data collection phases of the USAID/Morocco Reading for Success-Small Scale Experimentation activity, Summer Enrichment Program successful.

We would also like to thank Director of the Curricula, Fouad Chafiqi and his team for their continuous support in the implementation of the Reading for Success-Small Scale Experimentation activity.

Finally, we extend our appreciation to the regional and provincial offices that provided necessary field support through the mobilization of school inspectors, who effectively contributed to the development of the Early Grade Reading Assessment subtasks.

The following persons played a major role in the process:

Selma Talha Jebril	(Chemonics International)
Shamineh Byramji	(Chemonics International)
Rida Baidas	(Chemonics International)
Jordene Hale	(Chemonics International)
Louise Bahry	(School-to-School International)
Kayla Nachtsheim	(School-to-School International)
Beth Odenwald	(School-to-School International)
Selene Rangel	(School-to-School International)
Michel Rousseau	(School-to-School International)

CONTENTS

Acronyms	6
Executive Summary	7
Chapter 1: Introduction	9
1.1 Reading for Success-Small Scale Experimentation	9
1.2 Summary of the Summer Enrichment Program	10
1.3 Evaluation Design	11
1.4 What is EGRA?.....	12
Chapter 2: Methodology	13
2.1 School and Student Sample.....	13
2.2 EGRA Tool Description.....	14
2.3 Enumerator Training and Data Collection	15
2.4 Data Analysis	15
2.5 Limitations	16
Chapter 3: EGRA Results	17
3.1 EGRA Results by Subtask	17
3.2 EGRA Results by Gender	19
3.3 EGRA Results by Class Level	20
3.4 EGRA Results by Indicator.....	21
3.5 Relationships between the Results of EGRA Subtasks.....	22
Chapter 4: Contextual Variables and EGRA Performance	23
4.1 Student Questionnaire	23
Chapter 5: Conclusions and recommendations	26
Conclusion	26
Lessons Learned and Recommendations	27
Annexes	29
ANNEX A: Reliability of the EGRA Subtask Items	29
ANNEX B: EGRA Results – Additional Data.....	30
ANNEX C: Student Survey	32

LIST OF TABLES

Table 1: Participating Civil Society Organizations.....	9
Table 2: Number of Students in the Summer Reading Program.....	11
Table 3: Number of Students with at least 70% Attendance Rate	11
Table 4. Student Sample by Delegation and Treatment Group	13
Table 5. Student Sample by Gender and Treatment Group	14
Table 6. EGRA Results by Subtask and Treatment Group.....	17
Table 7. Percentage of Students with Improved Passage Reading Subtask Scores at Endline.....	21
Table 8. Percentage of Students with Improved Reading Comprehension Subtask Scores at Endline	22
Table 9. Correlation Matrix between the Number of Correct Answers for EGRA Subtasks at Endline	22
Table 10. Performance Gains by Parental Literacy and Treatment Group	24

LIST OF FIGURES

Figure 1: Key Stages of the Summer Program	10
Figure 2. Average Gains by EGRA Subtask and Treatment Group.....	18
Figure 3. Percentage of Zero Scores at Endline by EGRA Subtask and Treatment Group	18
Figure 4. Average Gains by EGRA Subtask, Student Gender, and Treatment Group.....	19
Figure 5. Percentage of Zero Scores at Endline by EGRA Subtask, Student Gender, and Treatment Group	19
Figure 6. Average Gains by EGRA Subtask, Class Level, and Treatment Group.....	20
Figure 7. Percentage of Zero Scores at Endline by EGRA Subtask, Class Level, and Treatment Group	21
Figure 8. Performance Gains by EGRA Subtask, Treatment Group, and Home Language	24

ACRONYMS

CSOs	Civil Society Organizations
EGRA	Early Grade Reading Assessment
MOEVT	Ministry of Education and Vocational Training
RFS-SSE	Reading for Success-Small Scale Experimentation
STS	School-to-School International
USAID	United States Agency for International Development

EXECUTIVE SUMMARY



This report presents key findings of the Morocco Reading for Success - Small Scale Experimentation (RFS-SSE) Early Grade Reading Assessment (EGRA) of the summer enrichment program. As part of the RFS-SSE project, 10 civil society organizations (CSOs) were awarded grants to implement summer reading enrichment activities during the summer of 2017. The summer activities complemented the RFS-SSE reading activities taught during the regular school year and encouraged students in grades 1 and 2 to continue developing their reading skills during summer vacation.

To evaluate the impact of the summer program on students' learning loss, RFS-SSE conducted a longitudinal study with a pretest-posttest design with an experimental and control group of students. Students were assessed at the end of the 2016-17 academic year, before summer activities, and then again at the start of the new academic year, after all summer program activities had completed. Students' performance was analyzed to determine whether children who participated in the summer enrichment program experienced less learning loss than children who had not participated. This report summarizes students' scores from the Baseline data collection in May and June 2017 and the Endline data collection in September 2017 and examines the changes in scores between those time points.

Both the experimental and control groups had statistically significant differences in average student scores from Baseline to Endline on four of the six subtasks: syllable identification, nonword reading, passage reading, and reading comprehension. However, there were no statistically significant differences between the treatment groups' change in scores on any subtask. That is, while there were statistically significant differences over time within each group, the differences in subtask scores from Baseline to Endline were not significantly different from one treatment group to the other.

While there were no statistically significant differences between the gains for the two groups, scores decreased for the three fluency subtasks and increased for three other subtasks, including reading comprehension. The decreases in scores for the syllable identification, nonword reading, and passage reading subtasks were statistically significant for both groups, as was the increase in scores for the reading comprehension subtask.

Two demographic characteristics were found to have a statistically significant correlation with students' scores: language spoken at home and parental literacy. Generally, statistical results showed that students

who reported speaking Amazigh at home had greater losses across subtasks than the students who reported speaking Darija at home, regardless of their treatment group. In regard to parental literacy, student gains were statistically significantly different according to parental literacy status, regardless of treatment group, on the reading comprehension subtask only. Students who indicated that their father is literate showed the greatest gains on the reading comprehension subtask from Baseline to Endline, regardless of if they were in the experimental group or the control group.

In general, student scores significantly decreased across fluency-based subtasks from Baseline to Endline data collection for both the experimental and control groups. However, non-fluency and comprehension subtasks showed some increases for many students. Most notably, the reading comprehension subtask results showed gains, sometimes statistically significant, regardless of class level, student gender, or treatment group.

CHAPTER 1: INTRODUCTION

1.1 Reading for Success-Small Scale Experimentation

The Reading for Success-Small Scale Experimentation (RFS-SSE) activity, a component of a broader United States Agency for International Development (USAID) initiative, was designed to reflect ongoing collaborations between USAID/Morocco and the Ministry of Education and Vocational Training (MOEVT) to improve reading instruction in Morocco. Conceived as a learning activity, the RFS-SSE project developed an evidence base of effective approaches that improve reading skills of students in targeted primary schools. The RFS-SSE project began when the MOEVT developed a 15-year education reform, called Vision 2030, and a set of medium-term activities for the period 2015–20. Reform efforts addressed a key weakness in the Moroccan educational system: poor reading skills among students at the primary level. The RFS-SSE project helped to inform revisions to the existing curriculum by providing data and evidence to support the envisioned changes.

This report will focus on one specific component of the RFS-SSE activity: a summer enrichment program. To reduce learning loss during the summer months, the RFS-SSE project awarded grants to 10 civil society organizations (CSOs) to implement summer reading enrichment activities during the summer of 2017. The list of participating CSOs is provided in Table 1. The summer activities complemented the RFS-SSE reading activities taught during the regular school year and encouraged students in grades 1 and 2 to continue developing their reading skills during a two-week period over the summer.

The selected CSOs participated in a series of capacity building activities in order to ensure administrative and technical implementation of their proposed programs. Each CSO selected RFS-SSE schools from seven (of the eight)¹ target provinces to participate in the summer program. The 10 grantees implemented and successfully completed their summer programs between the months of July and August. The reading enrichment activities engaged a total of 597 students (280 girls and 317 boys).

TABLE 1: PARTICIPATING CIVIL SOCIETY ORGANIZATIONS

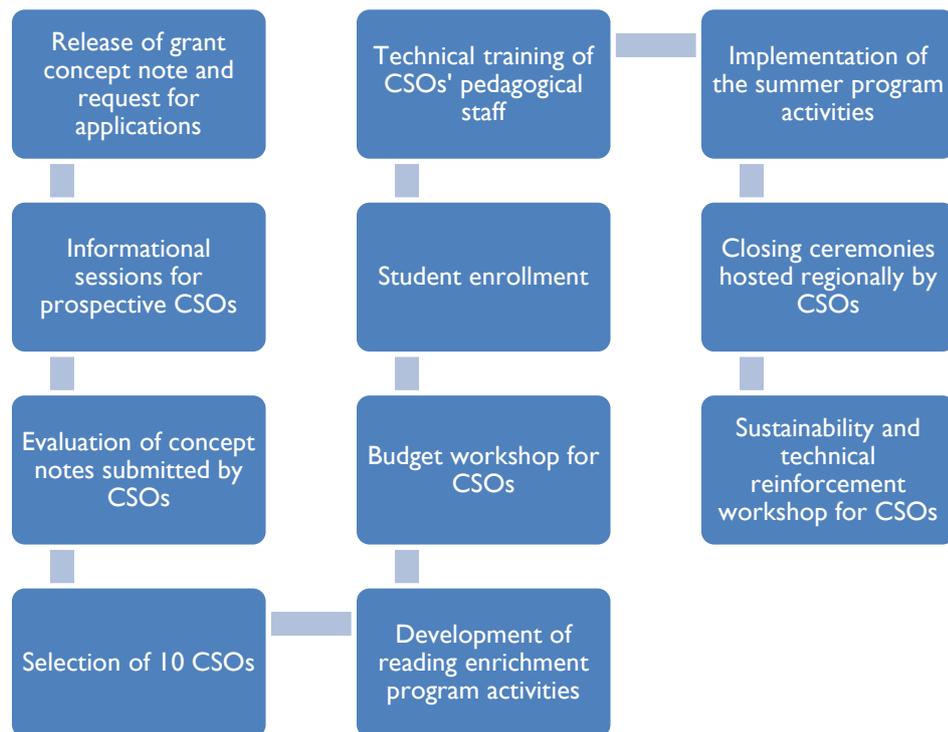
Province	Name of the organization	Year of creation
Oujda	Association de Coopération pour le Développement et la Culture	2014
Oujda	Union nationale des femmes des Maroc Oujda	1969
Inzegane	Scoutisme Hassania Maroc	1993
Inzegane	Association Moultaqa des jeunes pour le développement	2008
Tiznit	Association Forum de création et communication	2007
Taounate	Association Chraga pour le développement et la communication	2003
El Hajeb	Association Chantier Jeunesse Maroc (CJM) section Alhajeb	2015
Temara	Association Marocaine pour l'Education de la Jeunesse - section Temara	1976
Temara	Association Marocaine des Petits Débrouillards (AMPD)	2005
Kenitra	Association Sanad Elkhair	2012

¹ RFS-SSE initially planned to award grants for summer programs in all eight of the RFS-SSE target provinces. However, at the time of CSO selection, there was only one application from the Figuig province, and, unfortunately, the proposal score did not meet the minimum requirements necessary to be accepted into the grants program.

1.2 Summary of the Summer Enrichment Program

Summer reading programs are a new concept in Morocco; therefore, RFS-SSE conducted a series of capacity building trainings for the CSOs to prepare them to effectively implement enrichment activities that fostered a culture of reading outside the classroom. The training series brought together participants from each CSO and included topics such as program development, budget development, and technical reinforcement (see Figure 1). The main technical training last three days and focused on relating different activities, such as art, dance, theater, and song, to reading; imagining role-playing as a teaching and learning technique; and designing reading-based tools for use with students.

FIGURE 1: KEY STAGES OF THE SUMMER PROGRAM



Each of the 10 summer programs took place in either July or August 2017, and each program ran for an average of two weeks. Each CSO was allowed the flexibility to be creative and innovative with the design and implementation of their program; however, each program focused on engaging students in interactive student-centered reading games that would boost their reading skills and their confidence. These extracurricular reading activities took place in schools, public places, and community settings, which also encouraged an increased awareness and family involvement. Parents and community leaders committed to support the activities as well as advocated for the necessary institutional support from local officials.

RFS-SSE monitored the total number of students participating in the summer enrichment programs as well as the percentage of students attending at least 70 percent of the activities.

Table 2 presents the total number of students who participated in the summer enrichment programs, disaggregated by grade level and gender.

TABLE 2: NUMBER OF STUDENTS IN THE SUMMER READING PROGRAM

	Girls n	Boys n	Total n
Grade 1	141	150	291
Grade 2	139	167	306
Total	280	317	597

Table 3 presents the number of students who attended at least 70 percent of the CSOs’ programmed summer activities, disaggregated by grade level and gender. Daily sign-in sheets monitored students’ attendance. The attendance rate was calculated based on the number of days each student attended divided by the total number of days CSOs conducted activities.

TABLE 3: NUMBER OF STUDENTS WITH AT LEAST 70% ATTENDANCE RATE

	Girls n	Boys n	Total n
Grade 1	120	128	248
Grade 2	125	147	272
Total	245	275	520

1.3 Evaluation Design²

The RFS-SSE project examined whether summer enrichment activities related to reading could help reduce learning loss when students are out of school for summer break. To answer this question, the RFS-SSE project team conducted a longitudinal study with a pretest-posttest design with an experimental and control group of students. Students were assessed using the Early Grade Reading Assessment (EGRA) in May/June 2017³ at the end of the academic year and prior to the start of summer enrichment activities. These same students in both the experimental and control groups were assessed again in September 2017. Students in the experimental group participated in summer enrichment activities during the intervening months while students in the control group did not participate in such activities.

Students’ performance was analyzed to determine whether children who participated in the summer enrichment program experienced less learning loss than children who had not participated. This report summarizes students’ scores from the Baseline data collection in May and June 2017 and the Endline data collection in September 2017; it places a focus on changes in scores between those time points.

² This report only includes results from the study of the RFS-SSE project summer enrichment program. Throughout this report, “Baseline” refers to the data collected for this study prior to the start of the summer enrichment program in May and June 2017; “Endline” refers to the data collected for this study after the end of the summer enrichment program in September 2017. Prior to this summer study, the RFS-SSE project also completed a study of the effectiveness of the activities that were delivered during the normal academic year. Those results can be found in the report “USAID/Morocco: Reading for Success-Small Scale Experimentation Activity (RFS-SSE) Early Grade Reading Assessment Endline Report” prepared by School-to-School International and Chemonics International, Inc. in September 2017.

³ 455 students were assessed in May and 397 students were assessed in June. Because the CSOs partnered with schools that had participated in the RFS-SSE intervention, some of the students who participated in the CSOs’ summer programs had been assessed in May 2017 as part of the RFS-SSE internal impact evaluation. Therefore, students who had already taken the EGRA in May were not asked to re-take it in June. Because students did not receive any significant amount of instruction between May and June, all student results from May and June were combined to create one summer baseline dataset.

1.4 What is EGRA?

EGRA was developed in consultation with cognitive scientists, early grade reading experts, research methodologists, and assessment specialists with funding assistance from USAID, the World Bank, and other international donors. This tool assesses student performance on the basic foundational skills required for fluency in reading by measuring the skills needed for reading acquisition. Although many students are not yet fluent readers in the early grades, the EGRA allows researchers to capture what all students, even the “nonreaders,” can do and determine where they are in the developmental path towards becoming fluent readers. The EGRA does not provide individual student scores or measure teacher effectiveness. Generally, it takes several years before the full effect of improved curriculum and teacher training is fully realized in the EGRA.

Early Grade Reading Assessment (EGRA)

- Six subtasks that measure literacy skills
- Administered in one-to-one format outside of class
- Lasts 20 minutes and developed in Modern Standard Arabic

EGRA is an individually administered, oral assessment that requires approximately 20 minutes per student. The test is administered to one student at a time by a trained enumerator—not teachers—in a location outside of the classroom. The enumerator begins by explaining the assessment to the student and asking if the student agrees to participate. Consent is always optional and no student is required to take the assessment. The enumerator creates a relaxed environment for the student and assures the student that the assessment is not used for a grade. The enumerator then begins by asking the student questions aloud and having the student respond aloud. For specific subtasks, the enumerator places a paper stimulus in front of the student containing letters or words, and the enumerator asks the student specific questions about the stimuli. The Summer EGRA administered consisted of the same six subtasks used to assess student performance during the school year, and which were selected during the tool development workshop in December 2015.

CHAPTER 2: METHODOLOGY

This chapter describes the RFS-SSE project summer enrichment program sampling, data collection, analysis of results, and study limitations.

2.1 School and Student Sample

School Sample

Each of the 10 CSOs partnered with local schools to deliver summer enrichment program to students entering grade 2 or grade 3. This study collected data at all schools that participated in the summer enrichment program. Experimental schools were selected based on the CSO's location and existing relationship with the school. All CSOs partnered with schools that had participated in the RFS-SSE intervention during the 2016-17 academic year. Control schools were selected, by delegation, based on having similar attributes as the experimental school in that delegation. Therefore, for each delegation, a matched pair of schools was created: one experimental and one control. The final sample included 22 schools—10 control and 12 experimental—from four Académie Régionale d'Éducation et de Formation.⁴

Prior to the end of the academic year in June, each CSO coordinated with its paired school to enroll students in the summer program. The CSOs facilitated informational sessions with the School Directors, question and answer sessions with parents, and involved the teachers from the school. The specific start date of each summer program varied, but all programs started following the end of the school year in late June. Student enrollment took place in early to mid-June in order to collect the Baseline data prior to the end of the academic year.

To enhance comparability over time, the same students from the same schools were assessed at Baseline and Endline.⁵

Student Sample

Each delegation had a similar proportion of students in the control group as in the experimental group. In total, 49.25 percent of the students across all delegations were a part of the control group⁶, and 50.75 percent were a part of the experimental group (see Table 4).

TABLE 4. STUDENT SAMPLE BY DELEGATION AND TREATMENT GROUP

Académie Régionale d'Éducation et de Formation / Region	Delegation / Province	Control group n (%)	Experimental group n (%)	Total n (%)
Sous Massa	Inezgane-Aït Melloul	69 (20.97%)	70 (20.65%)	139 (20.81%)
	Tiznit	29 (8.81%)	31 (9.14%)	60 (8.98%)
Rabat-Kénitra	Témara	68 (20.67%)	70 (20.65%)	138 (20.66%)

⁴ Most CSOs partnered with one local school; however, CSOs in the Tiznit and Kenitra delegations each partnered with two schools in order to meet the minimum required students for the program.

⁵ 184 students were lost to follow up at Endline. At Baseline, 852 students were assessed; at Endline 676 students (79.34% of Baseline) were able to be assessed again. The results included in this report are only for those students for whom we have both Baseline and Endline data (i.e. 668 students).

⁶ Nine of the 329 control students reported having participated in another summer program outside of RFS-SSE. The difference in Endline scores across all EGRA subtasks for these 9 students was not significantly different from the remaining sample of control students. As a result, their results were not excluded from analysis.

	Kenitra	36 (10.94%)	35 (10.32%)	71 (10.63%)
Orient	Oujda-Angad	70 (21.28%)	67 (19.76%)	137 (20.51%)
Fès Meknès	Taounate	33 (10.03%)	37 (10.91%)	70 (10.48%)
	Elhajeb	24 (7.29%)	29 (8.55%)	53 (7.93%)
Total		329 (49.25%)	339 (50.75%)	668 (100.00%)

Table 5 provides additional detail, by delegation, of the students' gender and treatment group, showing nearly equal numbers across gender and treatment group at Endline.

TABLE 5. STUDENT SAMPLE BY GENDER AND TREATMENT GROUP

	Girls n (%)	Boys n (%)	Total n (%)
Control group	164 (49.85%)	165 (50.15%)	329 (49.25%)
Experimental group	176 (51.92%)	163 (48.08%)	339 (50.75%)
Total	340 (50.90%)	328 (49.10%)	668 (100.00%)

2.2 EGRA Tool Description

Two comparable EGRA tools were developed at the start of the RFS-SSE project.⁷ One version was administered at Baseline, and the other version was administered at Endline.⁸ The EGRAs developed for this study consisted of six subtasks: phonemic awareness, syllable identification, nonword reading, passage reading, reading comprehension, and listening comprehension.⁹

Both versions of the EGRA consist of the same six subtasks. The first three subtasks—phonemic awareness, syllable identification, and nonword reading—contain the same items, but the order of the items was rearranged within each subtask. The stories and related comprehension questions for the last three subtasks—passage reading, reading comprehension, and listening comprehension—are different but comparable in each of the two versions of the assessment.¹⁰

Several stories were developed and piloted during the original tool development process in 2015. RFS-SSE piloted four stories for the passage reading subtask and three stories for the listening comprehension subtask. After piloting all stories, the results were analyzed, and the two stories for each subtask which had the most similar results were selected for the operational assessments. This analysis included comparing the descriptive statistics, zero scores, cumulative frequency, and correlation with the first three subtasks for each story that was piloted.¹¹

⁷ A detailed explanation of the tool development process is presented in “USAID/Morocco: Reading for Success-Small Scale Experimentation Activity (RFS-SSE) Early Grade Reading Assessment Baseline Report” prepared by School-to-School International and Chemonics International, Inc. in May 2016.

⁸ Both of the EGRA tools were also administered during the larger RFS-SSE project's study conducted between January 2016 and May 2017. One tool was administered in January 2016, September 2016, and September 2017 (Endline for the summer enrichment program); the second version was administered in May 2016 and May 2017 (Baseline for the summer enrichment program).

⁹ A detailed explanation of the tool content is presented in “USAID/Morocco: Reading for Success-Small Scale Experimentation Activity (RFS-SSE) Early Grade Reading Assessment Endline Report” prepared by School-to-School International and Chemonics International, Inc. in September 2017.

¹⁰ A detailed explanation of the comparability of the two tools is presented in “USAID/Morocco: Reading for Success-Small Scale Experimentation Activity (RFS-SSE) Early Grade Reading Assessment Endline Report” prepared by School-to-School International and Chemonics International, Inc. in September 2017.

¹¹ Additionally, the mean difference in scores between the two stories was found to be less than 1/10th of a standard deviation; therefore, per the *Early Grade Reading Assessment (EGRA) Toolkit: Second Edition*, statistical equating was not required. For the passage reading/reading comprehension stories, the mean difference on the reading passage scores was 1.35 (1/10 SD=1.6). The mean difference on the reading comprehension scores was 0.032 (1/10 SD=0.036). For the listening

Some of the subtasks are timed, with students having up to one minute to complete the subtask, in order to measure the speed with which students are able to identify graphemes (syllable identification), decode grapheme-phoneme correspondence (nonword reading), and read connected text with fluency and accuracy (passage reading and reading comprehension). Accuracy is captured on all subtasks.

In addition to the EGRA, a short survey was administered to each student following completion of the assessment. The survey asked questions about student and family demographics and teacher practices. The student survey was developed at the same time as the EGRA in late 2015 by the working group convened in Rabat, Morocco.¹²

2.3 Enumerator Training and Data Collection

Enumerator Training

The MOEVT selected enumerators who were equally sourced from each of the eight delegations of the RFS-SSE activity. All enumerators were chosen by the National Center for Evaluations and Exams, a part of the MOEVT, and were either orientation inspectors or orientation advisors and counselors with certifications and field experience in evaluation.

All enumerators who collected data for this study had previously participated as enumerators in an earlier study in the RFS-SSE project. The RFS-SS project team facilitated a three-day refresher training prior to the start of the both the Baseline and Endline summer data collection. During the refresher trainings in May and September 2017, the enumerators were re-trained to administer all six EGRA subtasks on tablets using the electronic data capture application Tangerine. Enumerator performance was regularly monitored throughout training and data collection by local supervisors and international EGRA experts through frequent check-ins and daily data monitoring. Interrater reliability was not conducted during data collection due to the extremely high levels of enumerator agreement during prior data collections. During the data collection in May 2016, enumerator consistency during training ranged from 96.33 percent to 100 percent agreement depending on the subtask, and interrater reliability during data collection ranged from 97.72 percent to 100.00 percent depending on the subtask.

Data Collection

Baseline data collection occurred between May 8–24, 2017 and June 11–14, 2017; Endline data collection occurred from September 18–26, 2017. To ensure the quality of data collection, supervisors oversaw each data collection team and daily discussed progress and encountered problems with their teams. Supervisors verified that each enumerator reached as many of the participants as possible at each school and ensured that each enumerator completed a daily data collection report. Enumerator teams uploaded data from their tablets to a cloud database each evening of data collection. The STS EGRA coordinator reviewed and tallied data each night, discussing any concerns with the supervisors as necessary.

2.4 Data Analysis

Timed and untimed subtask scores were calculated using the data collected at Baseline and Endline. Since the same students were assessed at both time points, STS's researchers were able to make comparisons of subtask scores between Baseline and Endline. By comparing the mean difference scores

comprehension stories, the mean difference in scores was 0.004 (1/10 SD=0.029).

¹² A detailed explanation of the tool development process is presented in "USAID/Morocco: Reading for Success-Small Scale Experimentation Activity (RFS-SSE) Early Grade Reading Assessment Baseline Report" prepared by School-to-School International and Chemonics International, Inc. in May 2016.

of students in the control and experimental groups, researchers can measure the impact of the treatment—the summer enrichment program—on students. Additionally, STS correlated data from the student survey with subtask scores to identify demographic characteristics linked to EGRA performance.

2.5 Limitations

Short period of exposure to intervention

On average, the CSOs conducted activities with students during a two-week period over the summer. This means that each child received only 10 days, with an average of 4 hours a day, of educational support between the end of the academic year in June 2017 and the start of the new academic year in September 2017. The abbreviated nature of the summer program is a limitation to the intervention that should be considered when reviewing results.

Lack of Institutional Review Board (IRB)

An institutional review board (IRB) is responsible for determining the acceptability of proposed research regarding institutional commitments and regulations, applicable laws, standards of professional conduct and practice, and ethical and societal norms. IRBs examine subject recruitment procedures, proposed remuneration, and the informed consent process. IRBs also evaluate the potential risks and benefits to participants, as outlined in each protocol.

The RFS-SSE team, in consultation with in-country partners, determined that there was no appropriate local IRB process for Morocco. To address this, RFS-SSE provided the MOEVT with details about the project’s research design and obtained permission from the MOEVT to conduct the study.

Reliability of self-report survey data

The students were asked to respond to several survey questions after completing the assessment. The survey asked about socio-economic indicators, common practices and behaviors in the classroom, and the home environment. The risk of response bias is inherent to self-reported data. With self-reported data, respondents may be inclined to provide responses that they think will be considered the “correct” answer or responses that they think are socially acceptable, even if inaccurate. Therefore, the self-reported survey data is a limitation to the results of this study.

Lack of qualitative data

This small study collected only EGRA and survey data from students. Qualitative data such as teacher or parent interviews was not collected. Given the short time span of the project intervention, the addition of qualitative data may have provided additional insights into the contextual factors that affected the success of the RFS-SSE summer program in improving reading outcomes. For future studies, particularly of this size and length, it is recommended that a qualitative data collection and analysis component be added.

CHAPTER 3: EGRA RESULTS

Detailed EGRA results of the summer Endline include average scores by subtask for all students; average scores by subtask disaggregated by treatment group, student gender, and student grade; and the proportion of students unable to answer a single item correctly by subtask.

To show how students' EGRA scores changed during the summer months, results are presented using gain scores, calculated as the difference in student scores from the Baseline assessment in May/June 2017 to the Endline assessment in September 2017. As gain scores rely on two-time points, only students who had scores at both Baseline and Endline were included in this analysis.

3.1 EGRA Results by Subtask

Table 6 summarizes the EGRA results of all students from the Baseline assessment and the Endline assessment. Table 3 summarizes the mean scores of the experimental and control groups for each of the six EGRA subtasks.

TABLE 6. EGRA RESULTS BY SUBTASK AND TREATMENT GROUP

EGRA subtask	Experimental group		Control group		Experimental group	Control group
	Baseline	Endline	Baseline	Endline	Change since Baseline	
Phonemic awareness	6.29	6.55	7.18	7.26	0.27	0.08
Syllable identification	51.50	44.14	54.66	47.57	-7.36*	-7.08*
Nonword reading	16.09	12.77	17.32	13.63	-3.33*	-3.69*
Passage reading	22.72	16.61	24.95	18.72	-6.12*	-6.23*
Reading comprehension	0.86	1.60	0.91	1.74	0.73*	0.83*
Listening comprehension	2.30	2.33	2.29	2.38	0.04	0.09

An asterisk (*) indicates that the average score difference between Baseline and Endline is statistically significantly different at $p < 0.05$.
Note: Experimental Group $n = 339$; Control Group $n = 329$.

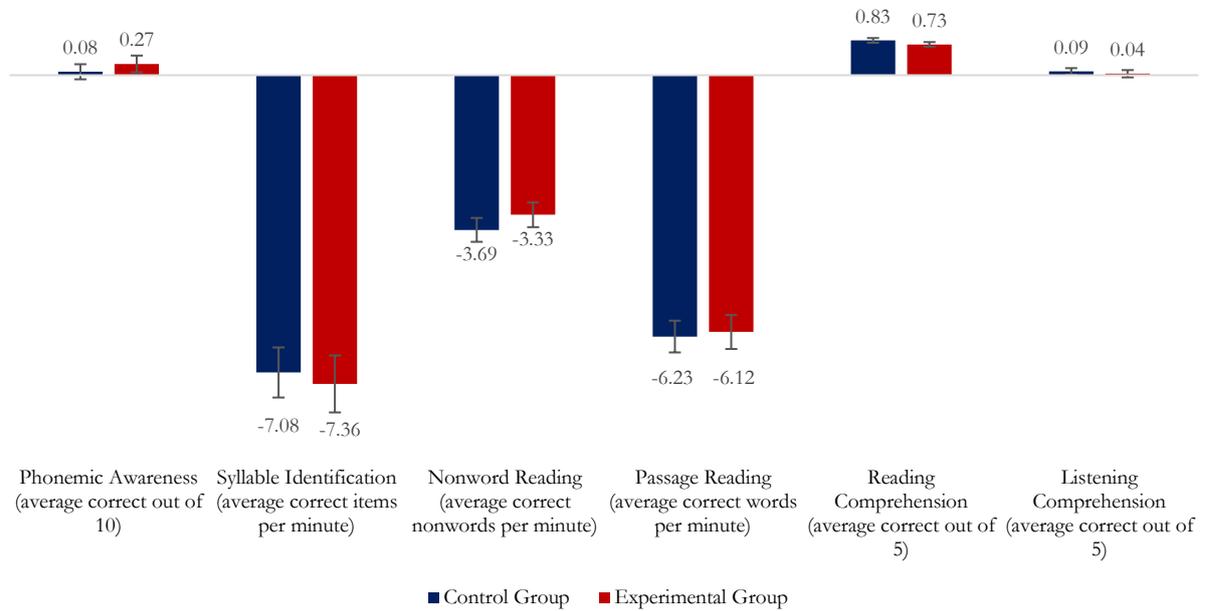
Both the experimental and control groups had statistically significant differences from Baseline to Endline on four of the six subtasks showed: syllable identification, nonword reading, passage reading, and reading comprehension. However, there were no statistically significant differences between the treatment groups' change in scores on any subtask. That is, while there were statistically significant differences over time within each group, the differences in subtask scores from Baseline to Endline were not significantly different from one treatment group to the other.

Figure 2 shows the gain scores by subtask and treatment group. Gain scores were computed by subtracting Baseline scores from Endline scores. Therefore, a positive gain score indicates an increase in the score from Baseline to Endline, whereas a negative gain score indicates a decrease in score from Baseline to Endline. While there were no statistically significant differences between the gains for the two groups, scores decreased for the three fluency subtasks and increased for three other subtasks, including reading comprehension. The decreases in

There were statistically significant differences over time within each group, but the differences in scores from Baseline to Endline were not significantly different between the experimental and control groups.

scores for the syllable identification, nonword reading, and passage reading subtasks were statistically significant for both groups, as was the increase in scores for the reading comprehension subtask.

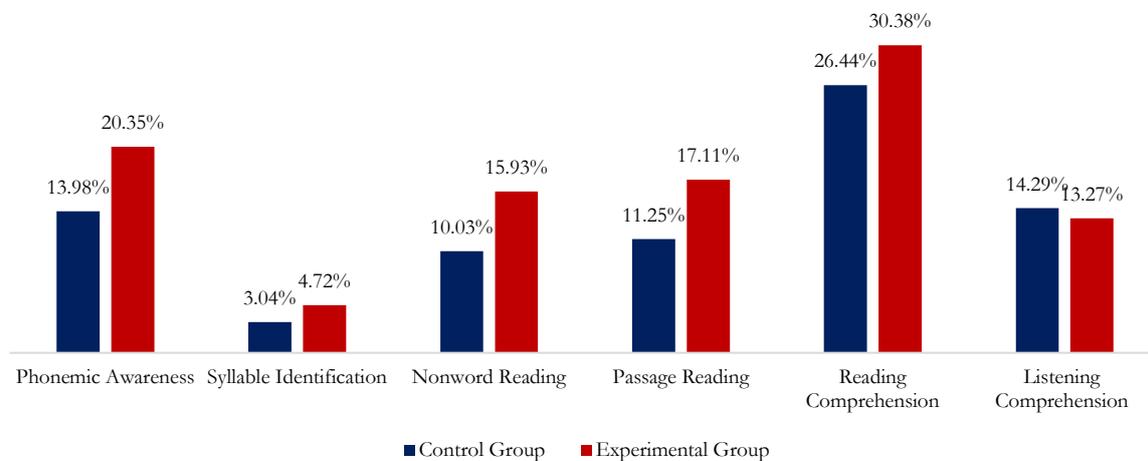
FIGURE 2. AVERAGE GAINS BY EGRA SUBTASK AND TREATMENT GROUP



Note: Control Group n = 329; Experimental Group n = 339.

Figure 3 shows the percentage of students, by group, who were unable to answer a single item correctly—or received a zero score—on each subtask at Endline. There were no statistically significant differences across groups in the proportion of zero scores. However, analysis shows a general trend: the proportion of students receiving zero scores is slightly higher for the experimental group than for the control group. The same trend is true at Baseline.

FIGURE 3. PERCENTAGE OF ZERO SCORES AT ENDLINE BY EGRA SUBTASK AND TREATMENT GROUP

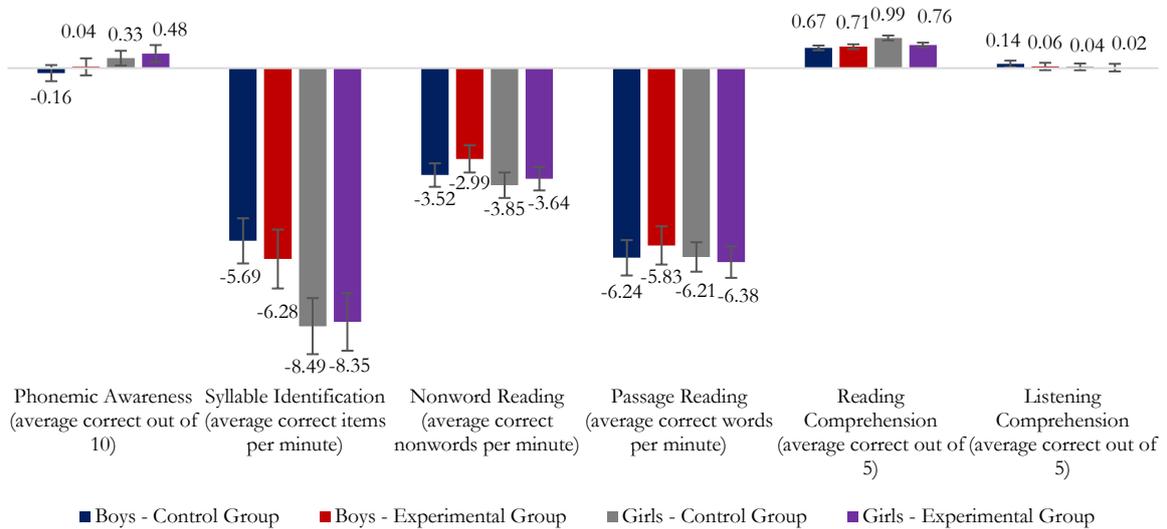


Note: Control Group n = 329; Experimental Group n = 339.

3.2 EGRA Results by Gender

Figure 4 shows the gain scores by subtask, treatment group, and student gender. There are no statistically significant differences by gender.

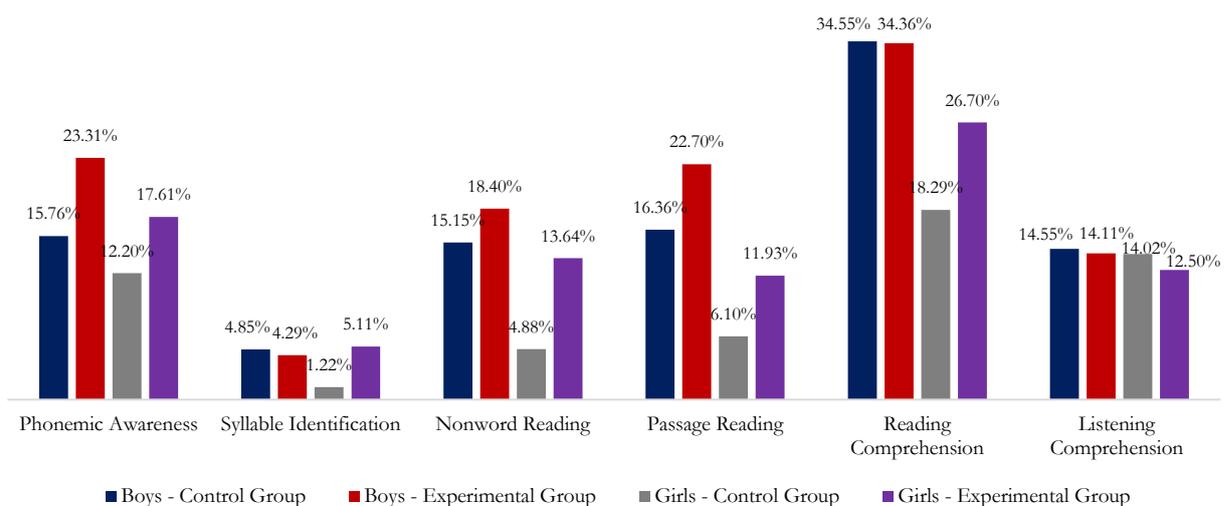
FIGURE 4. AVERAGE GAINS BY EGRA SUBTASK, STUDENT GENDER, AND TREATMENT GROUP



Note: Boys Control Group n = 165; Boys Experimental Group n = 163; Girls Control Group n = 164; Girls Experimental Group n = 176.

Figure 5 shows the percentage of students receiving zero scores on each subtask at Endline by treatment group and student gender. There were no statistically significant differences by gender.

FIGURE 5. PERCENTAGE OF ZERO SCORES AT ENDLINE BY EGRA SUBTASK, STUDENT GENDER, AND TREATMENT GROUP



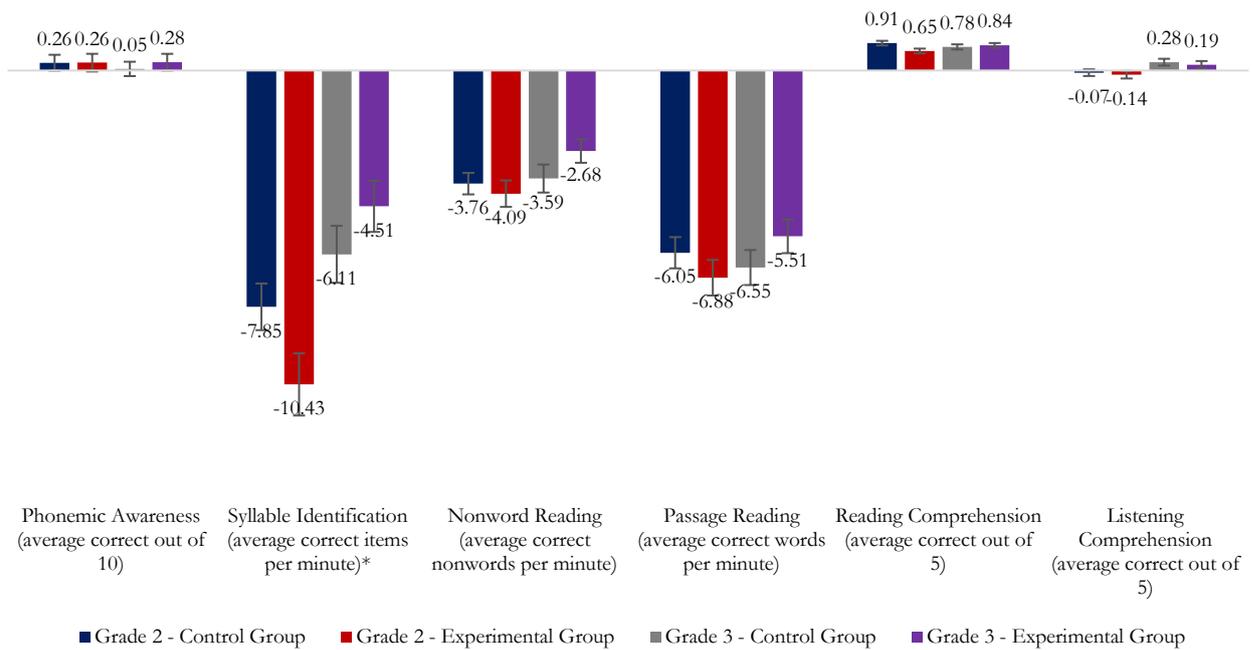
Note: Boys Control Group n = 165; Boys Experimental Group n = 163; Girls Control Group n = 164; Girls Experimental Group n = 176.

3.3 EGRA Results by Class Level

Figure 6 shows the gain scores by subtask, treatment group, and class level—grade 2 or grade 3.¹³ There was one statistically significant interaction between class level and treatment group. On the syllable identification subtask, grade 2 students in the experimental group saw statistically significant greater losses in fluency from Baseline to Endline than grade 3 students in the experimental group.

On the syllable identification subtask, grade 2 students in the experimental group saw statistically significant greater losses in fluency from Baseline to Endline than grade 3 students in the experimental group.

FIGURE 6. AVERAGE GAINS BY EGRA SUBTASK, CLASS LEVEL, AND TREATMENT GROUP

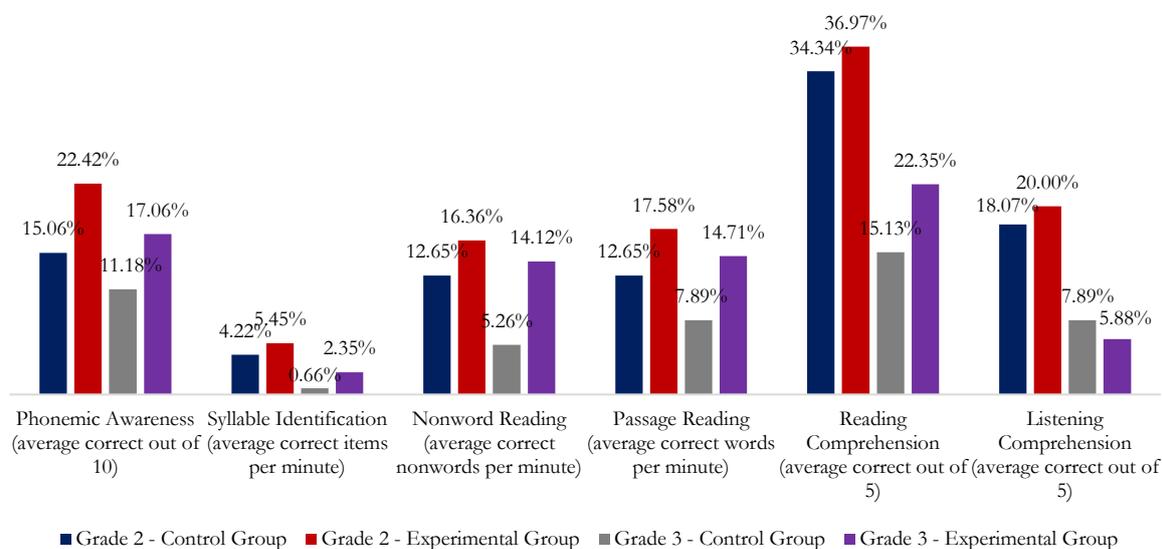


An asterisk (*) indicates a statistically significant interaction between class level and experimentation group at $p < 0.05$.
 Note: Grade 2 Control Group $n = 166$; Grade 2 Experimental Group $n = 165$; Grade 3 Control Group $n = 152$; Grade 3 Experimental Group $n = 170$.

Figure 7 shows the percentage of students receiving zero scores on each subtask at Endline, aggregated by class level and group. There were no statistically significant differences in the proportion of zero scores across groups.

¹³ Class level reflects the grade that students were entering at the end of the summer.

FIGURE 7. PERCENTAGE OF ZERO SCORES AT ENDLINE BY EGRA SUBTASK, CLASS LEVEL, AND TREATMENT GROUP



Note: Grade 2 Control Group n = 166; Grade 2 Experimental Group n = 165; Grade 3 Control Group n = 152; Grade 3 Experimental Group n = 170.

3.4 EGRA Results by Indicator

Improvement in students’ reading skills is measured as the percentage of students who scored higher at Endline than they had at Baseline when analyzing individual scores. As seen in Table 7, roughly 10 percent of male students (10.67 percent) and female students (10.29 percent) scored higher on the passage reading subtask at Endline than at Baseline. For both male and female students, there was no statistically significant difference between the performances of the control and experimental groups.

TABLE 7. PERCENTAGE OF STUDENTS WITH IMPROVED PASSAGE READING SUBTASK SCORES AT ENDLINE

Male students		
	n	%
Control group	19	11.52%
Experimental group	16	9.82%
Total	35	10.67%
Female students		
	n	%
Control group	17	10.37%
Experimental group	18	10.23%
Total	35	10.29%

Students’ improvement in reading comprehension was also analyzed. As seen in Table 8, nearly half (49.39 percent) of male students and a slightly higher proportion (58.82 percent) of female students scored higher on the reading comprehension subtask at Endline than they had at Baseline. There was no statistically significant difference between the control and experimental groups’ performances.

TABLE 8. PERCENTAGE OF STUDENTS WITH IMPROVED READING COMPREHENSION SUBTASK SCORES AT ENDLINE

Male students		
	n	%
Control group	77	46.67%
Experimental group	85	52.15%
Total	162	49.39%
Female students		
	n	%
Control group	104	63.41%
Experimental group	96	54.55%
Total	200	58.82%

3.5 Relationships between the Results of EGRA Subtasks

The relationships between EGRA subtasks results were analyzed to identify correlations between the scores on different subtasks and the EGRA overall. The possible correlation score ranges from zero to one. A correlation closer to one indicates a stronger relationship between how students performed on that particular subtask and how they performed on the EGRA overall. A correlation closer to zero indicates a weaker relationship between how students performed on that particular subtask and how they performed on the EGRA overall. Negative correlations indicate that as the score on the subtask increases, the score on the EGRA overall decreases.

All the correlations from the summer Endline EGRA are statistically significant ($p < 0.000$) and positive; this indicates that students who scored the highest on one subtask are more likely to have scored the highest on the other subtasks as well (see Table 9). Correlations were particularly strong between the syllable identification, nonword reading, and passage reading subtask. While this is not uncommon, among all the EGRA subtasks, scores on the listening comprehension subtask showed the weakest correlation to performance on the other subtasks.

TABLE 9. CORRELATION MATRIX BETWEEN THE NUMBER OF CORRECT ANSWERS FOR EGRA SUBTASKS AT ENDLINE

	Phonemic awareness	Syllable identification	Nonword reading	Passage reading	Reading comprehension	Listening comprehension
Phonemic awareness	1.0					
Syllable identification	0.51	1.0				
Nonword reading	0.49	0.85	1.0			
Passage reading	0.43	0.78	0.82	1.0		
Reading comprehension	0.44	0.75	0.75	0.84	1.0	
Listening comprehension	0.27	0.38	0.31	0.38	0.48	1.0

An asterisks (*) indicates that all correlations are statistically significant at $p < 0.000$.

CHAPTER 4: CONTEXTUAL VARIABLES AND EGRA PERFORMANCE

4.1 Student Questionnaire

Through the student questionnaire, enumerators asked students for familial and academic characteristics that may influence their learning. The RFS-SSE project team analyzed these variables in relationship to students' progress from the Baseline to the Endline on the EGRA subtasks. The aim was to identify which factors are linked to student progress in acquiring literacy skills. This report presents results presented only for those variables which were found to have a statistically significant relationship to EGRA results.

Language Spoken at Home

The EGRA used in this study was administered in Modern Standard Arabic which corresponds to language of instruction for all students participating in this study. However, students do not always speak the same language at home. Therefore, students were asked what language they speak at home.

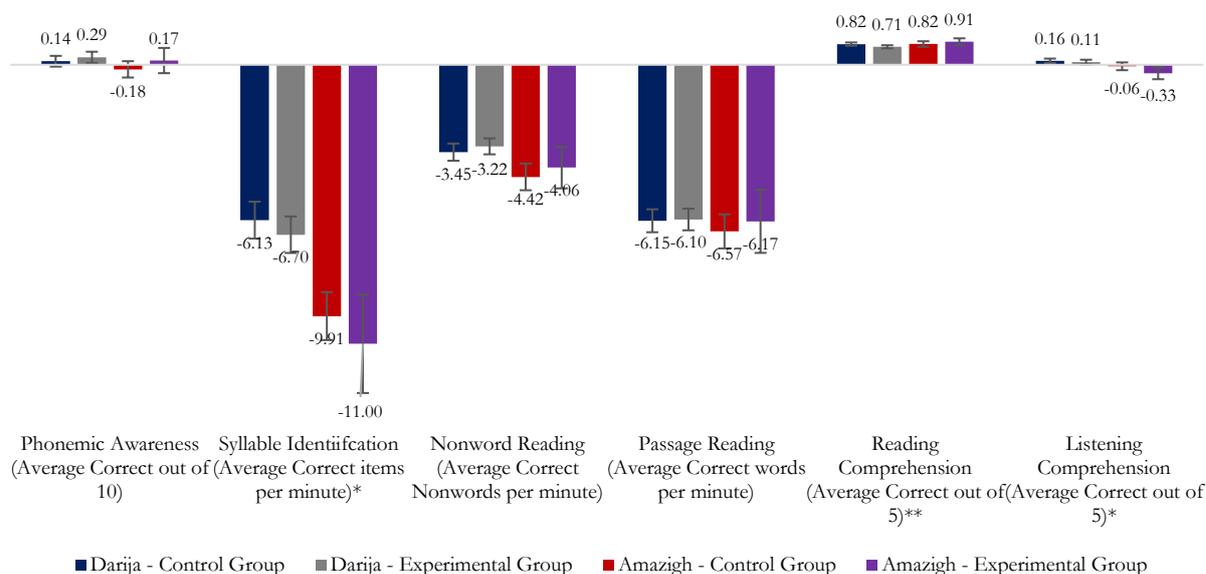
The progress of students varied for three of the six EGRA subtasks as a function of the language spoken at home and by their treatment group. Most (n=555, 83.08 percent) students surveyed speak Darija at home, while one-fifth (n=137, 20.51 percent) of the students surveyed speak Amazigh at home. Generally, statistical results showed that students who reported speaking Amazigh at home had greater losses across subtasks than the students who reported speaking Darija at home, regardless of their treatment group.

On the syllable identification and listening comprehension subtasks, students who reported speaking Amazigh at home had statistically significantly greater losses over the summer than students who reported speaking Darija at home, regardless of treatment group.

Results on the reading comprehension subtask indicate that students increased comprehension over the summer regardless of treatment group or home language. However, there was a statistically significant interaction between treatment group and home language for the reading comprehension subtask. Overall, students in the experimental group who spoke Amazigh at home showed the greatest gains, and more specifically, students in the control group who spoke either language made the second greatest gains on this task. Students in the experimental group who reported speaking Darija at home showed the least gains in Reading Comprehension.



FIGURE 8. PERFORMANCE GAINS BY EGRA SUBTASK, TREATMENT GROUP, AND HOME LANGUAGE



An asterisk (*) indicates that the gain score is statistically significantly different between reported home languages at $p < 0.05$.

Two asterisks (**) indicates a statistically significant interaction between home language and treatment groups at $p < 0.05$.

Note: Mean score estimates are noted adjacent to each bar in the figure. The confidence interval is indicated by the lines at the top of each bar. The confidence interval indicates a range of values that is likely to encompass the true value.

Note: Darija Control Group $n = 244$; Darija Experimental Group $n = 282$; Amazigh Control Group $n = 83$; Amazigh Experimental Group $n = 54$.

Status of Literacy at Home

On the reading comprehension subtask only, student gains were statistically significantly different according to parental literacy status, regardless of treatment group. That is, students who indicated that their father is literate showed the greatest gains on the reading comprehension subtask from Baseline to Endline, regardless of if they were in the experimental group or the control group. For both the control and experimental groups, scores were generally higher for students coming from homes where only the father could read than those of students who come from homes where only the mother could read. Table 10 shows the mean score on each subtask for the control and experimental groups by parental literacy.



TABLE 10. PERFORMANCE GAINS BY PARENTAL LITERACY AND TREATMENT GROUP

	Control group, parental literacy							
	Neither parent (n=53)		Mother only (n=61)		Father only (n=17)		Both parents (n=198)	
	Mean score	Standard deviation	Mean score	Standard deviation	Mean score	Standard deviation	Mean score	Standard deviation
Phonemic awareness	0.40	2.87	0.38	3.23	0.71	2.78	-0.15	3.40
Syllable identification	-4.32	13.30	-8.70	9.79	-9.99	9.33	-7.08	10.41
Nonword reading	-4.04	4.38	-3.75	5.09	-5.00	4.33	-3.46	5.45
Passage	-5.49	6.05	-5.87	7.84	-8.06	5.55	-6.38	6.86

	Control group, parental literacy							
	Neither parent (n=53)		Mother only (n=61)		Father only (n=17)		Both parents (n=198)	
	Mean score	Standard deviation	Mean score	Standard deviation	Mean score	Standard deviation	Mean score	Standard deviation
reading								
Reading comprehension*	0.68	0.92	1.08	1.08	0.94	1.09	0.78	1.01
Listening comprehension	0.02	1.32	0.28	1.37	0.41	1.28	0.03	1.44
	Experimental group, parental literacy							
	Neither parent (n=52)		Mother only (n=75)		Father only (n=40)		Both parents (n=172)	
	Mean score	Standard deviation	Mean score	Standard deviation	Mean score	Standard deviation	Mean score	Standard deviation
Phonemic awareness	0.40	3.73	0.80	4.13	0.20	4.87	0.01	3.09
Syllable identification	-6.65	15.79	-7.48	11.40	-6.88	14.55	-7.63	11.39
Nonword reading	-4.65	6.90	-2.71	4.82	-3.40	6.71	-3.18	4.74
Passage reading	-6.63	9.81	-5.88	6.84	-6.98	7.19	-5.86	7.02
Reading comprehension*	0.40	1.11	0.79	0.96	0.95	1.13	0.76	0.97
Listening comprehension	-0.06	1.56	0.19	1.54	0.00	1.48	0.01	1.66

An asterisk (*) indicates that the gain score is statistically significantly different depending on parental literacy at $p < 0.05$.

CHAPTER 5: CONCLUSIONS AND RECOMMENDATIONS

Conclusion

EGRA Results

In general, student scores decreased statistically significantly across fluency-based subtasks from Baseline to Endline data collection for both the experimental and control groups. However, non-fluency and comprehension subtasks showed small gains for many students. Most notably, the reading comprehension subtask results showed gains, sometimes statistically significant, regardless of class level, student gender, or treatment group. Generally, these results correspond to results from previous assessments of the RFS-SSE project in terms of gender differences and trends across contextual variables.

During the design stage of the summer programs, each local organization was requested to enroll at least 40 students in order to exceed a minimum number of students required for the EGRA sample size. Amongst the 10 selected organizations, in total 597 students participated in the summer programs; however, only 520 of these students attended 70 percent or more of the program. This high level of enrollment, in a short inscription period, demonstrates the interest and importance for such summer programs.

At the request of USAID, in order to have cohesive results between RFS-SSE components 1 and 2, it was agreed to use EGRA as the measurement tool for the summer activity. However, as outlined in this report, although there were statistically significant differences from Baseline to Endline for both groups (experimental and control), there was no statistically significant differences between the two groups. A variety of factors contributed to this outcome, most of which have been outlined below as lessons learnt and recommendations for future programs. However, the two most significant contributions to these factors are: (1) the length of the summer programs, as two-week programs were not long enough to make a lasting impact on specific reading subtasks that are measured via EGRA; and (2) the evaluation did not include any qualitative data which limits one's ability to explore the "why" behind the results .

Qualitative Feedback on the Summer Program

During the implementation, numerous field visits were completed allowing the RFS-SSE team and USAID to observe the CSOs' animators in action with the students and to ensure overall quality of implementation. It also afforded the team the opportunity to discuss the programs with students, parents, animators, and even community council members that were involved with the programs. These in-person discussions provided insights about the depth of the impact that these programs were having for some of the students. Indeed, based on feedback from the parents and animators, it was often mentioned that these programs led to a shift in the students' attitudes towards reading, and allowed them to have a positive and pleasurable experience with reading outside of the classroom. Through the different activities, there was a perceptible growth in students' confidence and interaction with other students. Additionally, the most remarkable change in parents was their realization of the importance of reading, and notably reading over the summer. This was best seen through the parents' event attendance and engagement in the program.

Similarly, through the various capacity building activities, trainings, and the implementation of these programs, the CSOs were exposed, often for the first time, to the importance of reading and the positive effects it can have on a child's development. Through these programs, the CSOs have become ambassadors within their communities for these types of summer reading programs.

RFS-SSE Program Lessons Learned and Recommendations

Length of Summer Program

The duration of future summer reading programs should be longer than two weeks. Based on RFS-SSE's experience, in order for the summer programs to have a significant effect on reducing reading loss during the summer months, the programs need to be longer than two weeks. This would require grant awards to include sufficient funds to allow for such a duration of programming.

CSO Technical Training Content

In order to be able to work with a variety of CSOs (small, large, new, established, etc.), it is necessary to sufficiently train the CSOs on both program design and technical pedagogy. It is best to provide several different types of trainings and at several points in time before the implementation of the summer programs. The CSOs should be equipped with the necessary tools to effectively design summer reading programs.

Additionally, it is important that the CSO participants who undergo training on the technical pedagogy are the same individuals who will be implementing the summer reading programs. If not sufficiently trained, an individual will not be fully equipped with the requisite skills to implement and facilitate a summer program that focuses on reading. Some areas of training could include, but are not limited to: (1) educational programming, (2) methods, strategies and tools to maintain and improve reading skills in early grades, and (3) criteria for selection or development of reading material (content, design, etc.).

Complementary Capacity Building Trainings for CSOs

In addition to technical trainings on the pedagogy of reading for the CSOs, at least a two-day communication and monitoring and evaluation (M&E) capacity building training should also be organized before implementation of the summer program in order to equip them with the skills necessary to monitor the effectiveness and implementation fidelity well.

Evaluation Tool

To best measure the impact of the summer programs, it is encouraged to develop an evaluation tool that can measure both the qualitative and quantitative aspects of the programs. For RFS-SSE, it was required to assess the summer programs using an EGRA tool, in hopes that the project could measure the retention of what was taught to students during the academic year under component 1 of the project. However, as the EGRA measures reading skill, and not curriculum content, this tool may not have been the most appropriate to measure the impact of the summer programs, especially given the short timeline of said programs.

Additionally, as it was most efficient to access the population of students to be assessed during the school year, RFS-SSE conducted the EGRA in a school setting. Thus, the Baseline was administered in schools in June 2017 and the Endline was administered in schools in September 2017, while the majority of the summer programs took place (for two weeks) during July 2017. Therefore, there was a gap between the end of the summer programs and the Endline EGRA administration, which did not allow RFS-SSE to test the impact of the intervention immediately following implementation. In order to have a more timely and specific measurement of impact of the summer program, it would be best to develop a pre- and post-questionnaire measuring reading skills that could be administered at the very beginning and end of the summer program.

Promote Reading Clubs All Year

Given the qualitative feedback gathered throughout the summer, there appears to be solid interest and positive engagement in programs such as this one. Therefore, in addition to summer reading programs, we suggest that schools promote and organize extra-curricular reading programs in schools during the school year to ensure continuity during the summer. This could be done through Reading Clubs or intermittent but regular initiatives, such as reading classes or class competitions with incentives for students.

Learning from the Community for Optimal Program Design

As focus groups are a great qualitative data generator that ensure community engagement and recommendations, we suggest organizing focus groups before the start of the summer programs, at the program design stage, and at the end of the programs to ensure that the recommendations from the community—mainly the parents and teachers—are taken into consideration as much as possible.

Promote the Summer Reading Programs

We also encourage the promotion of summer programs through social media and local media to highlight the efforts and leadership of the local organizations that will continue to raise awareness about the importance of reading and engagement. Additionally, by advertising the summer programs during the spring, it will ensure that people are aware of the programs and help ensure student enrollment and attendance.

ANNEXES

ANNEX A: Reliability of the EGRA Subtask Items

The analysis of the quality of the assessment consisted of analyzing the reliability of the full assessment using coefficient alpha with percent correct scores on each subtask. This index varies between 0 and 1, where a value closer to 1 indicates that the performance of the students can be easily generalized. This analysis treats each subtask as a single item with a score ranging from 0.0 to 1.0 for each student (their percent correct on the subtask). This method of measuring internal consistency may show lower estimates of alpha due to the small number of items being using in the calculation (equal to the number of subtasks administered and the low variability in scores for three of the six subtasks, as coefficient alpha is heavily dependent on the number of items used and the variability in responses to the items. In Tables 1 and 2 we see the results of this analysis. The overall EGRA internal consistency for Baseline is 0.82, and for Endline is slightly lower at 0.85.

TABLE 1. RELIABILITY MEASURES AT BASELINE USING PERCENT CORRECT

Reading subtasks	Number of students	Subtask-Other subtasks correlation	Alpha if subtask is deleted
Phonemic Awareness (percentage correct)	668	0.51	0.84
Syllable Identification (percentage correct)	668	0.74	0.77
Nonword Reading (percentage correct)	668	0.74	0.77
Passage reading (percentage correct)	668	0.80	0.75
Reading comprehension (percentage correct)	668	0.63	0.79
Listening comprehension (percentage correct)	668	0.39	0.84
Alpha			0.82

TABLE 2. RELIABILITY MEASURES AT ENDLINE USING PERCENT CORRECT

Reading subtasks	Number of students	Subtask-Other subtasks correlation	Alpha if subtask is deleted
Phonemic Awareness (percentage correct)	668	0.50	0.88
Syllable Identification (percentage correct)	668	0.81	0.80
Nonword Reading (percentage correct)	668	0.79	0.82
Passage reading (percentage correct)	668	0.80	0.80
Reading comprehension (percentage correct)	668	0.80	0.79
Listening comprehension (percentage correct)	668	0.43	0.87
Alpha			0.85

ANNEX B: EGRA Results – Additional Data

TABLE 1. AVERAGE GAIN SCORES FOR ALL EGRA SUBTASKS

	Control			Experimental		
	Count	Mean	SE	Count	Mean	SE
Phonemic Awareness (Average Number Correct out of 10)	329	0.08	0.18	339	0.27	0.20
Syllable Identification (Average Number of Correct Items Per Minute)	329	-7.08	0.60	339	-7.36	0.68
Nonword Reading (Average Number of Correct Nonwords Per Minute)	329	-3.69	0.28	339	-3.33	0.29
Passage Reading (Average Number of Correct Words Per Minute)	329	-6.23	0.38	339	-6.12	0.41
Reading Comprehension (Average Number Correct out of 5)	329	0.83	0.06	339	0.73	0.06
Listening Comprehension (Average Number Correct out of 5)	329	0.09	0.08	339	0.04	0.09

TABLE 2. PERCENT OF STUDENTS RECEIVING ZERO SCORES FOR ALL EGRA SUBTASKS AT BASELINE AND ENDLINE

	Baseline				Endline			
	Control		Experimental		Control		Experimental	
	n	Mean	n	Mean	n	Mean	n	Mean
Phonemic Awareness (Average Number Correct out of 10)	329	13.68%	329	25.66%	329	13.98%	339	20.35%
Syllable Identification (Average Number of Correct Items Per Minute)	329	2.13%	329	3.54%	329	3.04%	339	4.72%
Nonword Reading (Average Number of Correct Nonwords Per Minute)	329	8.81%	329	13.57%	329	10.03%	339	15.93%
Passage Reading (Average Number of Correct Words Per Minute)	329	6.69%	329	9.44%	329	11.25%	339	17.11%
Reading Comprehension (Average Number Correct out of 5)	329	49.54%	329	49.85%	329	26.44%	339	30.38%
Listening Comprehension (Average Number Correct out of 5)	329	16.41%	329	15.34%	329	14.29%	339	13.27%

TABLE 3. AVERAGE GAIN SCORES FOR ALL EGRA SUBTASKS BY STUDENT GENDER AND GROUP

	Girls					
	Control			Experimental		
	Count	Mean	SE	Count	Mean	SE
Phonemic Awareness (Average Number Correct out of 10)	164	0.33	0.24	176	0.48	0.28
Syllable Identification (Average Number of Correct Items Per Minute)	164	-8.49	0.92	176	-8.35	0.95
Nonword Reading (Average Number of Correct Nonwords Per Minute)	164	-3.85	0.42	176	-3.64	0.38

Passage Reading (Average Number of Correct Words Per Minute)	164	-6.21	0.48	176	-6.38	0.52
Reading Comprehension (Average Number Correct out of 5)	164	0.99	0.08	176	0.76	0.08
Listening Comprehension (Average Number Correct out of 5)	164	0.04	0.11	176	0.02	0.12
	Boys					
	Control			Experimental		
	Count	Mean	SE	Count	Mean	SE
Phonemic Awareness (Average Number Correct out of 10)	165	-0.16	0.27	163	0.04	0.28
Syllable Identification (Average Number of Correct Items Per Minute)	165	-5.69	0.74	163	-6.28	0.97
Nonword Reading (Average Number of Correct Nonwords Per Minute)	165	-3.52	0.39	163	-2.99	0.45
Passage Reading (Average Number of Correct Words Per Minute)	165	-6.24	0.58	163	-5.83	0.63
Reading Comprehension (Average Number Correct out of 5)	165	0.67	0.08	163	0.71	0.08
Listening Comprehension (Average Number Correct out of 5)	165	0.14	0.11	163	0.06	0.12

TABLE 4. AVERAGE GAIN SCORES FOR ALL EGRA SUBTASKS BY STUDENT GRADE AND GROUP

	Grade 2					
	Control			Experimental		
	Count	Mean	SE	Count	Mean	SE
Phonemic Awareness (Average Number Correct out of 10)	166	0.26	0.26	165	0.26	0.29
Syllable Identification (Average Number of Correct Items Per Minute)	166	-7.85	0.78	165	-10.43	1.03
Nonword Reading (Average Number of Correct Nonwords Per Minute)	166	-3.76	0.36	165	-4.09	0.44
Passage Reading (Average Number of Correct Words Per Minute)	166	-6.05	0.52	165	-6.88	0.59
Reading Comprehension (Average Number Correct out of 5)	166	0.91	0.08	165	0.65	0.08
Listening Comprehension (Average Number Correct out of 5)	166	-0.07	0.11	165	-0.14	0.13
	Grade 3					
	Control			Experimental		
	Count	Mean	SE	Count	Mean	SE
Phonemic Awareness (Average Number Correct out of 10)	152	0.05	0.24	170	0.28	0.28
Syllable Identification (Average Number of Correct Items Per Minute)	152	-6.11	0.95	170	-4.51	0.85
Nonword Reading (Average Number of Correct Nonwords Per Minute)	152	-3.59	0.47	170	-2.68	0.39
Passage Reading (Average Number of Correct Words Per Minute)	152	-6.55	0.58	170	-5.51	0.56
Reading Comprehension (Average Number Correct out of 5)	152	0.78	0.09	170	0.84	0.08
Listening Comprehension (Average Number Correct out of 5)	152	0.28	0.11	170	0.19	0.12

ANNEX C: Student Survey

بأية لغة تتحدث (ين) مع أسرتك بالمنزل؟	<input type="checkbox"/> الدارجة <input type="checkbox"/> العربية <input type="checkbox"/> الأمازيغية <input type="checkbox"/> الفرنسية <input type="checkbox"/> آخر <input type="checkbox"/> لا يعرف/ يرفض
ما هو مستواك الدراسي السنة الماضية؟	<input type="checkbox"/> التمهيدي <input type="checkbox"/> مستوى 1 <input type="checkbox"/> مستوى 2 <input type="checkbox"/> بدون <input type="checkbox"/> لا يعرف/ يرفض
بأية وسيلة تذهب (ين) إلى المدرسة غالبا؟ (ملاحظة: اختر إجابة واحدة)	<input type="checkbox"/> مشيا <input type="checkbox"/> بالدراجة <input type="checkbox"/> أستقل الحافلة <input type="checkbox"/> بالسيارة <input type="checkbox"/> وسيلة أخرى <input type="checkbox"/> لا يعرف/ يرفض
من يرافقك؟ (ملاحظة: اختر إجابة واحدة)	<input type="checkbox"/> وحدي <input type="checkbox"/> برفقة والداي / الوالدة <input type="checkbox"/> برفقة أختي / أخي <input type="checkbox"/> برفقة زملائي <input type="checkbox"/> برفقة شخص آخر <input type="checkbox"/> لا يعرف/ يرفض
<i>Cahier d'exercices</i>	
هل لديك دفتر التمارين؟	<input type="checkbox"/> نعم <input type="checkbox"/> لا <input type="checkbox"/> لا أعرف
هل يمكن أن أتصفحه؟	<input type="checkbox"/> نعم <input type="checkbox"/> لا
كم عدد الصفحات المملوءة؟	<input type="text"/>
كم عدد الصفحات المصححة في دفترك؟	<input type="text"/>

ما هي طبيعة التقييم؟	<input type="checkbox"/> نقطة عددية <input type="checkbox"/> تقدير <input type="checkbox"/> لا شيء
<i>Enseignant et famille</i>	
كيف يتصرف أو تتصرف معك الأستاذ (ة) عندما يكون جوابك صحيحا؟ (ملاحظة لا يجب على الممرر أن يقوم بقراءة الأسئلة للتلميذ)	<input type="checkbox"/> لا شيء <input type="checkbox"/> يثمن عملي <input type="checkbox"/> يعطيني هدية (لاصقة، قلم رصاص) <input type="checkbox"/> يعفيني من عمل أو واجب منزلي <input type="checkbox"/> لا يعرف/ يرفض <input type="checkbox"/> آخر شرح: _____
كيف يتصرف معك الأستاذ عندما يكون جوابك خاطئا؟ (ملاحظة لا يجب على الممرر أن يقوم بقراءة الأسئلة للتلميذ)	<input type="checkbox"/> لا شيء <input type="checkbox"/> يصحح <input type="checkbox"/> الأستاذ يوبخ التلميذ (ة) <input type="checkbox"/> لا يعرف/ يرفض <input type="checkbox"/> آخر شرح: _____
هل تتلقى مساعدة في إنجاز الواجبات المنزلية؟	<input type="checkbox"/> نعم <input type="checkbox"/> لا يعرف/ يرفض <input type="checkbox"/> لا
إذا كان الجواب نعم، من يساعدك؟	<input type="checkbox"/> أخ / أخت <input type="checkbox"/> الأم / الأب <input type="checkbox"/> الجد / الجدة <input type="checkbox"/> لا يعرف/ يرفض <input type="checkbox"/> شخص آخر شرح: _____
<i>Alimentation</i>	
هل أكلت (وجبة الفطور - وجبة الغداء) قبل المجيء إلى المدرسة؟	<input type="checkbox"/> نعم <input type="checkbox"/> لا <input type="checkbox"/> لا يعرف/ يرفض
هل تستفيد من الإطعام بالمدرسة؟	<input type="checkbox"/> نعم <input type="checkbox"/> لا <input type="checkbox"/> لا يعرف/ يرفض
من أين تحصل على وجبة الغداء؟	<input type="checkbox"/> في المنزل <input type="checkbox"/> أحمل معي مأكولات

	<input type="checkbox"/> أشتري في عين المكان <input type="checkbox"/> لا يعرف/ يرفض <input type="checkbox"/> آخر شرح: _____
<i>Absences</i>	
هل تغيبت عن المدرسة في الأسبوع الماضي؟	<input type="checkbox"/> نعم <input type="checkbox"/> لا <input type="checkbox"/> لا يعرف/ يرفض
إذا كان الجواب نعم، ماذا كان سبب غيابك؟	<input type="checkbox"/> لأنني كنت مريضا <input type="checkbox"/> لأنني استيقظت متأخرا <input type="checkbox"/> لأنني لم أكن أتوفر على أكل <input type="checkbox"/> بسبب يوم السوق أو الإعداد ليوم السوق <input type="checkbox"/> لأنه كان على الاعتناء بإخوتي <input type="checkbox"/> لأنه كان على الاعتناء بأحد أفراد عائلتي المريض <input type="checkbox"/> بسبب العمل بالمنزل <input type="checkbox"/> لم تكن هناك أي وسيلة للنقل <input type="checkbox"/> لأنني لم أكن أتوفر على البديلة المدرسية <input type="checkbox"/> لأنني أتلقى معاملة سيئة من طرف الأساتذة أو التلاميذ بالمدرسة <input type="checkbox"/> لأن المدرسة محفوفة بالخطر <input type="checkbox"/> لأنني أجد صعوبة في الدروس <input type="checkbox"/> لأن المدرسة غير مهمة <input type="checkbox"/> بسبب سوء حالة الجو <input type="checkbox"/> سبب آخر <input type="checkbox"/> لا يعرف/ يرفض
في الأسبوع الماضي وفي الصباح، هل وصلت متأخرا إلى المدرسة؟	<input type="checkbox"/> نعم <input type="checkbox"/> لا <input type="checkbox"/> لا يعرف/ يرفض
إذا كان الجواب نعم، لماذا؟	<input type="checkbox"/> لأنني كنت مريضا <input type="checkbox"/> لأنني استيقظت متأخرا <input type="checkbox"/> لأنه كان على الاعتناء بإخوتي <input type="checkbox"/> لأنه كان على الاعتناء بأحد أفراد عائلتي المريض

	<input type="checkbox"/> بسبب عمل آخر بالمنزل <input type="checkbox"/> لم تكن هناك أي وسيلة للنقل أو تأخر النقل <input type="checkbox"/> لأنني لم أجد بذلة مدرسية أو البذلة لم تكن جاهزة <input type="checkbox"/> لأنني أتلقى معاملة سيئة من طرف الأساتذة أو الأطفال بالمدرسة <input type="checkbox"/> بسبب سوء حالة الجو <input type="checkbox"/> سبب آخر <input type="checkbox"/> لا يعرف/ يرفض
<i>Lecture</i>	
من من أسرته يستطيع القراءة؟	<input type="checkbox"/> لا أحد <input type="checkbox"/> الأب <input type="checkbox"/> الأم <input type="checkbox"/> الأخ أو الأخت <input type="checkbox"/> آخر <input type="checkbox"/> لا يعرف/ يرفض
هل لديك كتاب اللغة العربية؟	<input type="checkbox"/> نعم <input type="checkbox"/> لا <input type="checkbox"/> لا يعرف/ يرفض
إذا كان الجواب نعم، هل تستعمله في المنزل؟	<input type="checkbox"/> نعم <input type="checkbox"/> لا <input type="checkbox"/> لا يعرف/ يرفض
إذا كان الجواب نعم، هل لديك كتب، مجلات، جرائد، بالمنزل؟	<input type="checkbox"/> نعم <input type="checkbox"/> لا <input type="checkbox"/> لا يعرف/ يرفض
إذا كان الجواب نعم، هل هناك شخص يقرأ معك الكتب عادة؟	<input type="checkbox"/> نعم <input type="checkbox"/> لا <input type="checkbox"/> لا يعرف/ يرفض
هل تتوفر عائلتك على (ملاحظة للممرر يمكن قبول : عدة أجوبة)	<input type="checkbox"/> مرحاض خارجي <input type="checkbox"/> مرحاض داخل البيت <input type="checkbox"/> حفرة <input type="checkbox"/> كهرباء <input type="checkbox"/> تلفاز <input type="checkbox"/> الثلجة

	<input type="checkbox"/> آلة غسيل <input type="checkbox"/> حاسوب <input type="checkbox"/> الانترنت <input type="checkbox"/> سيارة <input type="checkbox"/> لا يعرف/ يرفض
هل تتوفر عائلتك على الماء؟	<input type="checkbox"/> نعم <input type="checkbox"/> لا <input type="checkbox"/> لا يعرف/ يرفض
إذا كان الجواب نعم، مصدره:	<input type="checkbox"/> الصنبور <input type="checkbox"/> العين <input type="checkbox"/> خزان <input type="checkbox"/> بئر <input type="checkbox"/> آخر <input type="checkbox"/> لا يعرف/ يرفض
مع من تعيش بالمنزل؟	<input type="checkbox"/> الأم <input type="checkbox"/> الاب <input type="checkbox"/> الإخوة <input type="checkbox"/> الأجداد <input type="checkbox"/> آخر <input type="checkbox"/> لا يعرف/ يرفض
هل تقوم بأشغال خارج أوقات الدراسة؟	<input type="checkbox"/> نعم <input type="checkbox"/> لا <input type="checkbox"/> لا يعرف/ يرفض
إذا كان الجواب نعم، ما هي؟	<input type="checkbox"/> أشغال البيت <input type="checkbox"/> العمل في الحقل <input type="checkbox"/> شراء مستلزمات البيت <input type="checkbox"/> لا يعرف/ يرفض